REPRINTED FROM

PROCEEDINGS OF THE 2ND ANNUAL WILLIAM T. PECORA

MEMORIAL SYMPOSIUM

OCTOBER 25-29, 1976
SIOUX FALLS, SOUTH DAKOTA

INVENTORY AND MAPPING OF FLOOD INUNDATION USING INTERACTIVE DIGITAL IMAGE ANALYSIS TECHNIQUES

Wayne G. Rohde
Technicolor Graphic Services, Inc.
Sioux Falls, South Dakota

James V. Taranik
U.S. Geological Survey
Sioux Falls, South Dakota

Charles A. Nelson
Technicolor Graphic Services, Inc.
Sioux Falls, South Dakota

INVENTORY AND MAPPING OF FLOOD INUNDATION* USING INTERACTIVE DIGITAL IMAGE ANALYSIS TECHNIQUES

Wayne G. Rohde
Technicolor Graphic Services, Inc
Sioux Falls, South Dakota
James V. Taranik
U.S. Geological Survey
Charles A. Nelson
Technicolor Graphic Services, Inc.
Sioux Falls, South Dakota

Biography: Wayne G. Rohde received his B.S. and M.S. in forestry from the University of Michigan in 1969 and 1971, respectively. Mr. Rohde is Principal Forestry Applications Scientist at the U.S. Geological Survey's EROS Data Center. He has extensive experience in the application of aircraft and satellite sensing systems to agriculture and forestry resource management problems. Prior to joining the staff at the EROS Data Center, he was employed at Earth Satellite Corporation, U.S. Army Engineer Topographic Laboratories, and the University of Michigan, Willow Run Laboratories (now ERIM). Mr. Rohde is a member of the Society of American Foresters, Ecological Society of America, and American Society of Photogrammetry.

Biography: James V. Taranik received his B.Sc. and Ph.D. degrees in geology from Stanford University and Colorado School of Mines, respectively. As a commissioned officer on active duty with the U.S. Army, Jim served as Chief Geologist for the Corps of Engineers in Vietnam. In 1971 he joined the Iowa Geological Survey as Chief of Remote Sensing Laboratory. He was also appointed Adjunct Professor of Geology in the University of Iowa and Visiting Professor of Civil Engineering at Iowa State University, and in those capacities he offered courses in geological remote sensing at both institutions. Jim joined the U.S. Geological Survey in 1975 and serves as Principal Remote Sensing Scientist, in charge of geological applications at the Survey's EROS Data Center in Sioux Falls, South Dakota. Jim also serves as Adjunct Professor of Earth Sciences at the University of South Dakota. He is a member of the American Society of Photogrammetry.

Biography: Mr. Charles A. Nelson received his B.S. and M.S. degrees from the College of Forestry, University of Minnesota in 1969 and 1971, respectively. Mr. Nelson worked on the Forest Application Project at the NASA Johnson Space Center before becoming a Data Analyst in the Data Analysis Laboratory at the EROS Data Center. He is a member of the American Society of Photogrammetry.

ABSTRACT

Landsat digital data acquired in July 1975, were used to estimate the agricultural land in a 610,000 ha area inundated by the May - June 1975 flood in the Red River Valley in North Dakota. A controlled clustering technique was developed to derive training set statistics that could be used to classify the Landsat data with a maximum likelihood algorithm. Evaluation of initial results indicated misclassification between older residential areas and agricultural land which had been partially inundated. Fallow agricultural fields and fields recently plowed and planted were also misclassified as inundated.

*Presented at the Second Annual William T. Pecora Symposium, October 1976, Sioux Falls, South Dakota.

The Landsat digital data were stratified into three strata to minimize the variation in land cover and inundated agricultural land within each stratum. All strata boundaries were plotted on U.S. Geological Survey 1:24,000 scale maps and digitized with a sonic digitizer. Thirty-five control points were selected to develop a mapping transformation to register the digitized strata boundaries to the Landsat image with a mean residual registration error of less than 0.5 pixel.

A maximum likelihood algorithm was used to classify all picture elements within each stratum. After classification, the three strata were combined to reconstruct a classified image of the entire area. Evaluation of these results indicates that stratification of Landsat data prior to classification significantly reduces misclassification error.

The classified Landsat data were geometrically corrected with a first-order transformation and nearest-neighbor resampling technique to produce 1:24,000 scale map overlays with a mean residual error of approximately 0.5 pixel. The classification results were used in a multiphase sampling design to estimate that 379,167 ha (sampling error =4 percent) of agricultural land were affected by flood conditions.

This procedure provides a technique for rapid assessment of the extent of flood conditions on agricultural land and would provide a basis for designing a sampling framework to estimate the impact of flooding on crop production.

INTRODUCTION

Each year floods damage millions of acres of agricultural land in the United States. When floods occur, rapid appraisals of the location and extent of flood inundation are necessary to assess crop damages and impact on crop production.

Panchromatic aerial photographs have been utilized for many years to document flooding on agricultural lands. These photographs are usually acquired near the time of flood crest but there are serious operational problems. Inclement weather limits the times of data acquisition and flood crests often occur in hours of darkness. Because flood crests move slowly downstream, repeated overflights may be required. Also, silt-laden flood water is not easily discriminated from wet bare soil on panchromatic photographs. These problems limit the usefulness of panchromatic photographs for assessing flood damage.

Recent studies have shown that utilization of an infrared sensitive film allows assessment of the latent effects of flooding on soils and vegetation as much as one month after flood crest (Meyer and Welch, 1975, p. 1534-1542).

Hoyer, Hallberg, and Taranik (1974) indicate that a large difference in infrared reflectance between saturated soils covered partially by plant debris, and plants and soils are not affected by flooding can be recorded with remote sensing data acquired as much as 5 or 10 days after flood crest. McAdams (1976) described techniques for estimating flood produced crop damage over an area of approximately 540 ha using multidate color infrared photographs. When floods occur over extensive areas, however, the costs for acquiring and interpreting aerial photographs can be high.

Landsat imagery covers very large areas and, when analyzed properly, provides efficient estimates of resource parameters. Investigators have shown that when remote sensing imagery at several scales is available, it is often most efficient to first make a large number of fast, inexpensive measurements of a parameter (X_i), on small scale, low resolution imagery, and then correlate this parameter with the parameter of interest (Y_i), (Langley, 1969; Heller and Wear, 1969; Nichols, 1974; Gialdini and others, 1975). A second phase involves selecting a small number of sample units on which the parameter of interest (Y_i) is measured precisely on larger scale, higher resolution images. Statistical methods are then used to estimate the parameter of interest from the measurement made on each image type and to estimate the variability about the estimated total. Landsat data have been used in a multistage sampling framework of this type to estimate the acreage of irrigated agricultural land in Idaho (Packer, 1976).

The purpose of this study was to demonstrate the use of Landsat digital data and statistical sampling procedures to estimate the extent of flood inundation on agricultural land. The specific objectives were to estimate with accuracy of $\pm$ 10 percent at the 95 percent probability level, the area of agricultural land that was either partially or completely inundated.

STUDY AREA AND AVAILABILITY OF DATA

In May and June of 1975, the Red River Valley in North Dakota and Minnesota received precipitation about 250 mm above normal. During June 28-30, an area of about 16,000 km^2 received more than 100 mm of rainfall. An unofficial rain gage near Leonard, N. Dak. registered precipitation of 508 mm over the 3-day period.

Landsat data acquired over the area on July 14, 1975, indicate the extent of flooding in the Fargo-Moorhead areas. (figure 1) Several different types of flooding occurred: (1) overbank flooding caused by high stream stages along the lower reaches of the Sheyenne, Maple, and Red Rivers in North Dakota; (2) overbank flooding caused by constriction of streamflow by man-made structures and lodgment of debris in forested areas; and (3) flooding caused by accumulation of overland runoff behind levees, road and railroad embankments, and other manmade structures in both Minnesota and North Dakota. Damages exceeding $375,000,000 on approximately one million ha of agricultural land, with yield reductions ranging between 40 and 100 percent, were estimated by U.S. Department of Agriculture agencies for the Fargo-Moorhead area.

High-altitude (approximately 18,000 m) color infrared photographs were acquired on July 10, 1975 with the NASA RB-57 aircraft. The approximate scale of the photographs is 1:120,000. These photographs were used to locate training areas on the Landsat images for the digital classification (phase 1) and in the second phase of the multiphase sampling design.

DIGITAL IMAGE CLASSIFICATION

Landsat digital data for approximately 610,000 ha were analyzed using an interactive analysis procedure and a maximum likelihood classification algorithm. A controlled clustering technique, described by Fleming, Berkebile, and Hoffer (1975), was used to derive training set statistics for use with the classification algorithm.

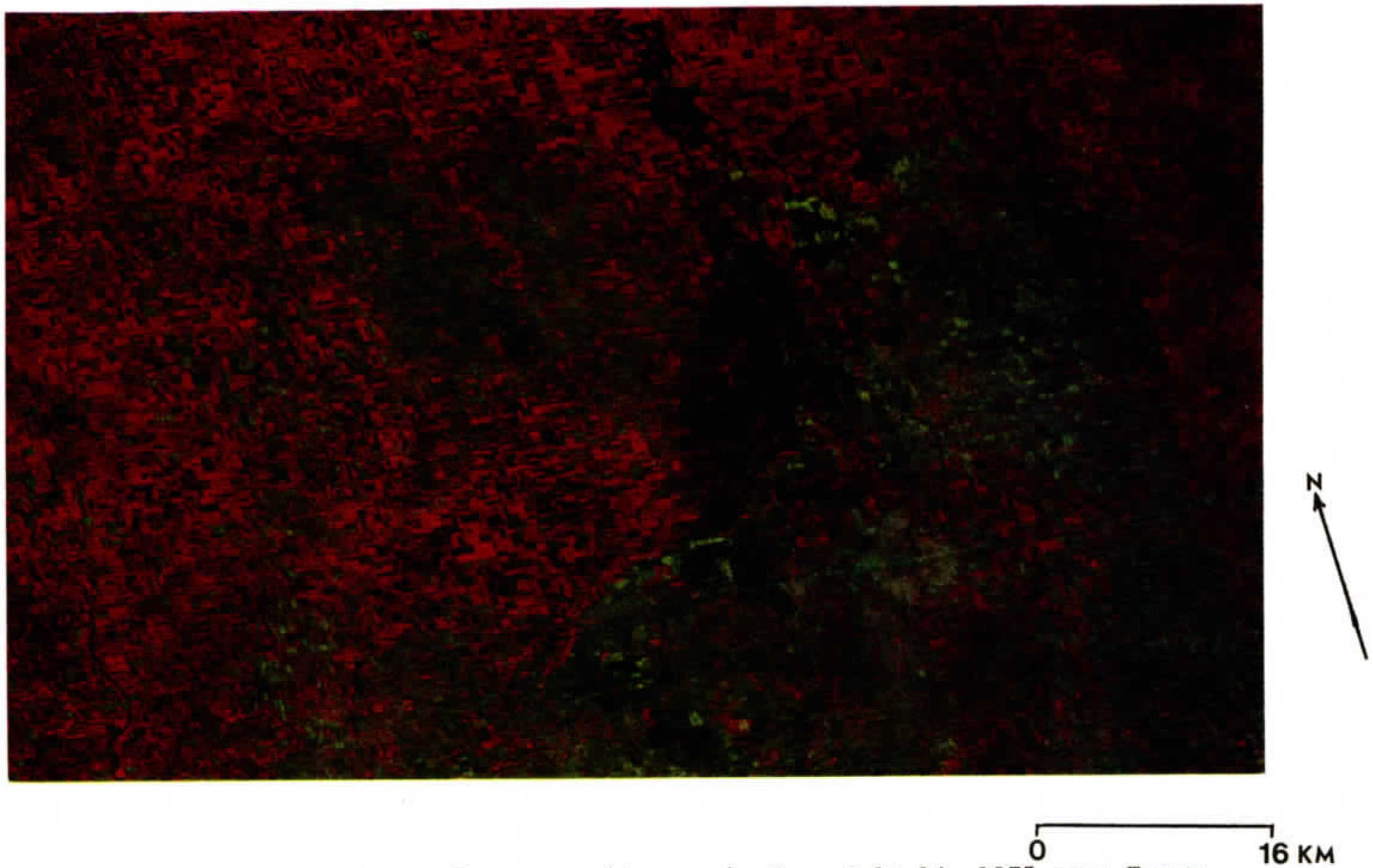

Figure 1.--Part of a Landsat color composite acquired on July 14, 1975 over Fargo, N. Dak. and Moorhead, Minn. area. Agricultural land affected by flooding appears in various shades of blue, gray, and blue-green. This image has not been geometrically corrected.

Five areas, each approximately 2,200 ha, were located on color infrared photographs (figure 2). Areas were selected which represent the range of flood conditions and land cover types throughout the area. The training areas were located on the Landsat data and brightness values of all pixels within the training areas were clustered into a large number of spectral classes. The clustering program provides a list of the mean, standard deviation, and covariance matrix for each spectral class and a line-printer map identifying the spectral class to which each pixel was assigned. The line-printer map of each training area was compared with a color infrared photograph. Each spectral class was then assigned to one of the ground cover classes shown in table 1. Each cover class was characterized by several spectral classes which were statistically similar. The statistics for all spectral classes were used to calculate a statistical measure of the separability between spectral classes in multidimensional space. The separability statistic and the interpreted cover class of each spectral class were used to determine which spectral classes could be combined. The final statistics file contained twenty-one spectral classes.

Table 1.--Land cover classification scheme used in mapping extent of flood conditions in the Fargo, N. Dak. - Moorhead, Minn. area in July, 1975.

<u>Cover type</u>

1. Urban
2. Agricultural land not affected by flood
3. Partially inundated - agricultural land partially affected by flood condition
4. Completely inundated - agricultural land where effects of flooding were apparent over entire field
5. Standing water
6. Wooded (both dry and flooded)

The training area statistics and Landsat data were put in a maximum likelihood classifier. The brightness value of each pixel in all four Landsat MSS bands and the mean and covariance matrix for each training class were used to calculate the likelihood that each pixel belonged to that training class. After all calculations were made, each pixel was assigned to the class it had the highest likelihood of belonging to. Evaluation of initial results indicated confusion between older residential areas and agricultural land which had been partly inundated. Fallow fields and fields recently plowed and planted were also misclassified as inundated.

The Landsat digital data were stratified into three strata to minimize the variation in land cover and inundated agricultural land within each stratum. A Landsat color composite, similar to that shown in figure 1, was enlarged to a scale of 1:250,000. All urban areas greater than approximately 65 ha were plotted on U.S. Geological Survey 1:24,000-scale maps, and digitized. They were then used to develop a transformation to register the digitized strata boundaries to the Landsat image with a mean residual error of less than 0.5 pixel (figure 3).

The strata boundaries were used to extract the image data associated with each strata. The original training statistics were modified to reduce the number of training classes and minimize the likelihood

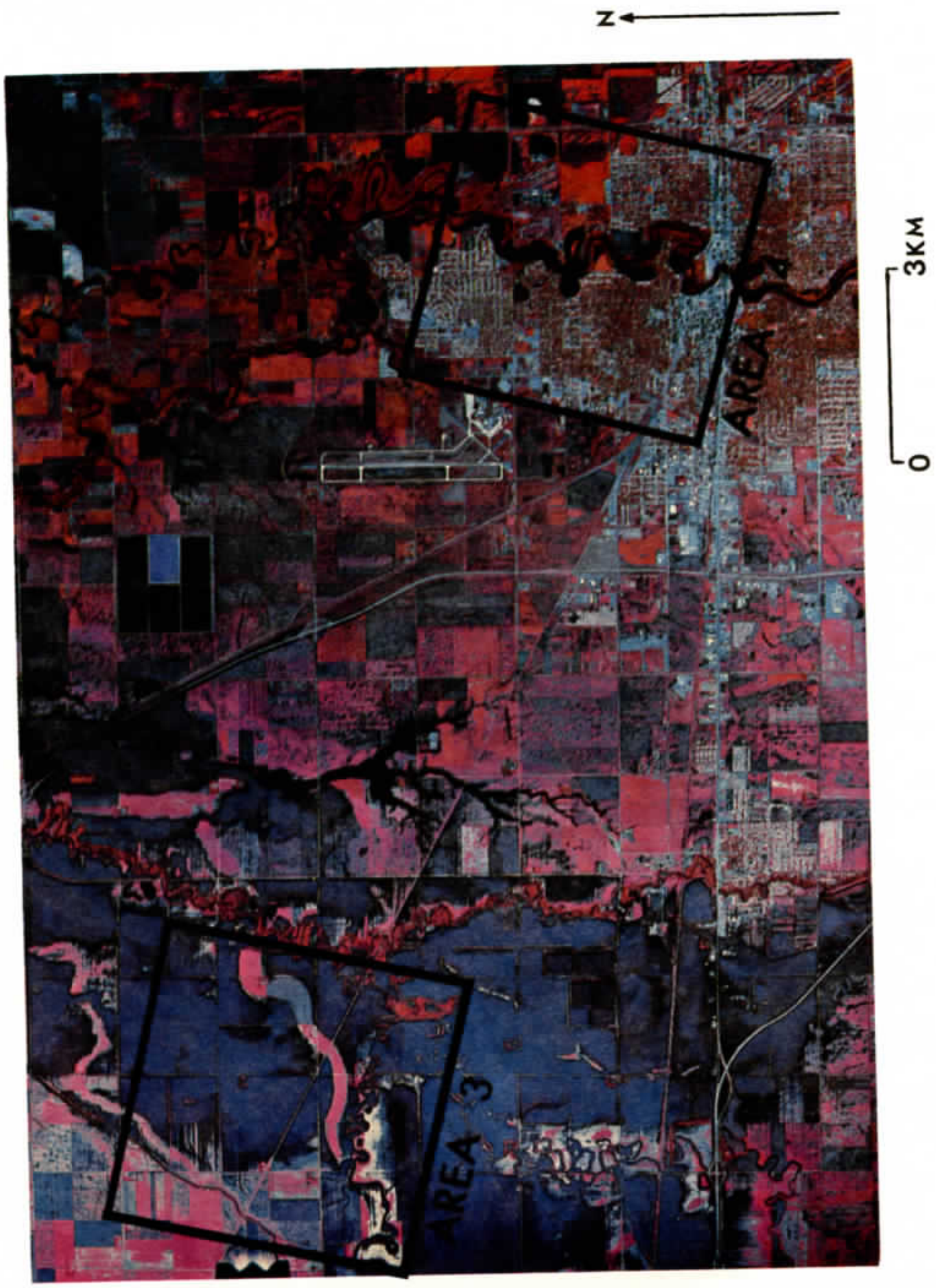

Figure 2.--Part of a NASA color infrared photograph taken on May 10, 1975, from an RB-57 aircraft at an altitude of 18,000m. Areas 3 and 4 are approximately 2,200ha. and their geographic locations were registered with Landsat data. All picture elements within each area were clustered into spectral classes for later input to a maximum likelihood classifier.

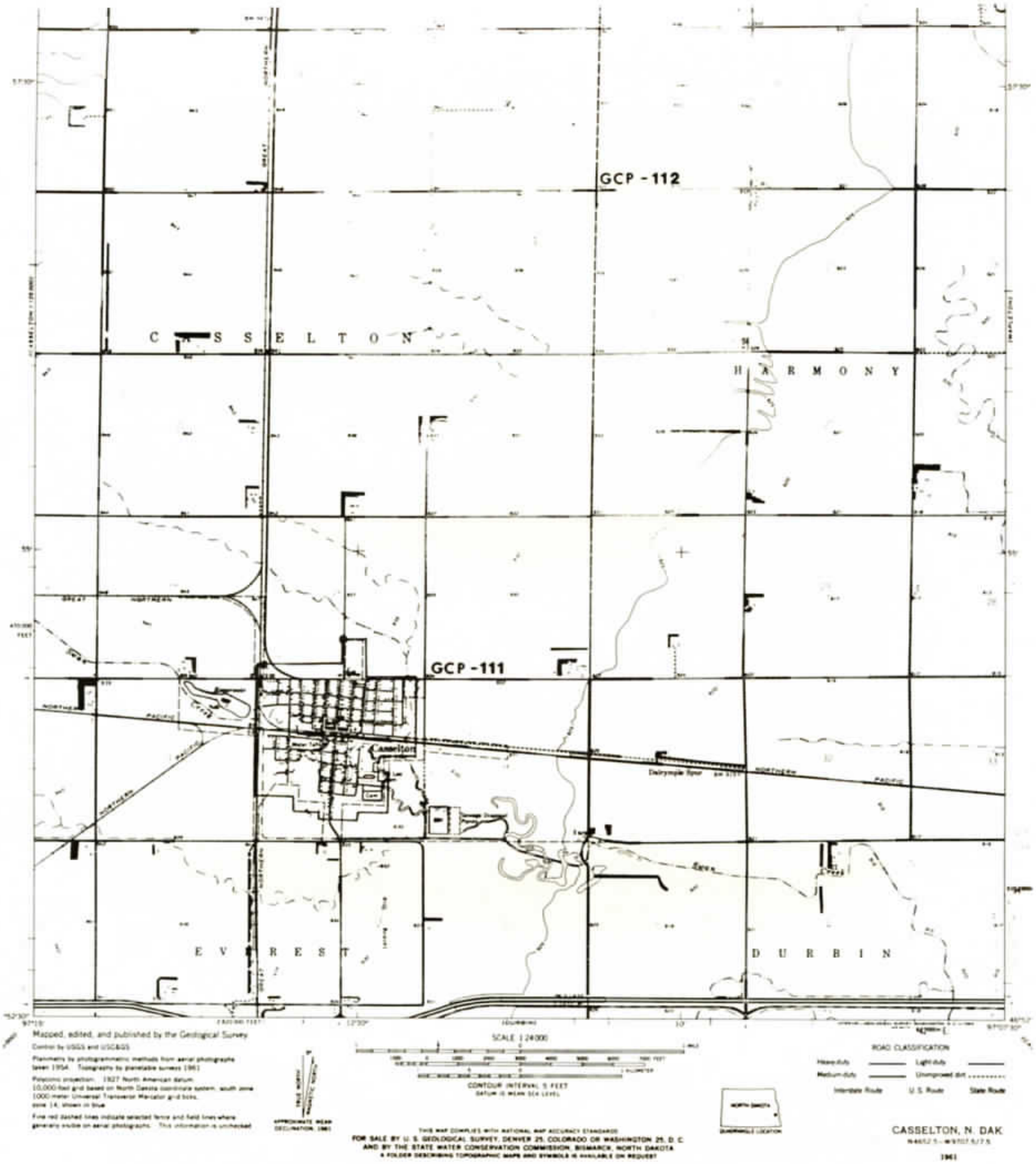

Figure 3.--A part of a photographically reduced topgraphic map. GCP-112 and GCP-111 are ground control points which were digitized with a sonic digitizer. The urban area of Casselton was outlined and digitized, thus, control points were used to develop a mapping transformation to register the digitized strata boundaries to the Landsat data.

of misclassification noted in the initial results. For example, within the urban stratum, all training classes associated with the agricultural lands were deleted from the statistics file since these cover classes did not exist in the urban areas. Similarly, within the agricultural land stratum, the urban training classes were deleted. A separate training statistics file was created for each stratum. The training statistics file and appropriate Landsat data for each stratum were input to a maximum likelihood classification algorithm to classify all picture elements. After classification, the three strata were added to reconstruct the original image. The final classification results are shown in figure 4. Evaluation of these results indicates that stratification of Landsat data prior to classification significantly reduces misclassification error. The Landsat classified data then were geometrically corrected with a first order mapping transformation and nearest neighbor resampling technique to produce 1:24,000 scale map overlays with a mean residual error of approximately 0.5 pixel.

The classification results provide an indication of the areal extent of the effect of flooding on agricultural land. Because the data are in digital format, the area within each cover class can be easily tabulated. However, this provides no indication of classification accuracy nor can a confidence statement be placed on the tabulated figures. In order to estimate the area within each cover class and to calculate a confidence interval for each estimate, a multiphase sampling procedure utilizing the Landsat digital classification results in the first phase was used.

SAMPLING PROCEDURE FOR AREA ESTIMATES

A two-phase sampling scheme was used to estimate the area of agricultural land affected by flood conditions. In two-phase sampling, sample units are the same size in each phase. The Landsat digital classification results provided the quick estimates in the first phase. A subsample of the first-phase sample units was selected for more precise measurement in the second phase. A least squares regression of estimates from the second phase on estimates from the first phase was used to adjust estimates made from the first phase.

Before allocating sample units to either the first or second phase, the size of a specific sample unit and number of sample units required must be determined. The optimum size of a primary sample unit (PSU) can be estimated by observing changes in the coefficient of variation for PSU's of varying sizes. In this project, four different sized PSU's were used: 100 pixels, 400 pixels, 1,600 pixels, and 3,600 pixels. The entire project area was gridded into the PSU's described above. The area of flooded cropland was calculated for each PSU. The mean, standard deviation, and coefficient of variation for the different PSU sizes were calculated (table 2).

The optimum size of a PSU can be chosen by selecting a sample unit size which minimizes the coefficient of variation and the size of the sample unit. In this project, a PSU of 40x40 pixels (1,600 pixels) was selected as an optimum size. Assuming a PSU size of 1,600 pixels, the number of sample units required for measurement at each phase was estimated, based on an expected correlation between measurements made at each phase, the cost ratio of obtaining measurements at each phase, and the desired accuracy level. To achieve an estimate with an accuracy of $\pm$ 10 percent at the .95 probability level, it was estimated that 200 first-phase samples and 30 second-phase samples would have to be measured.

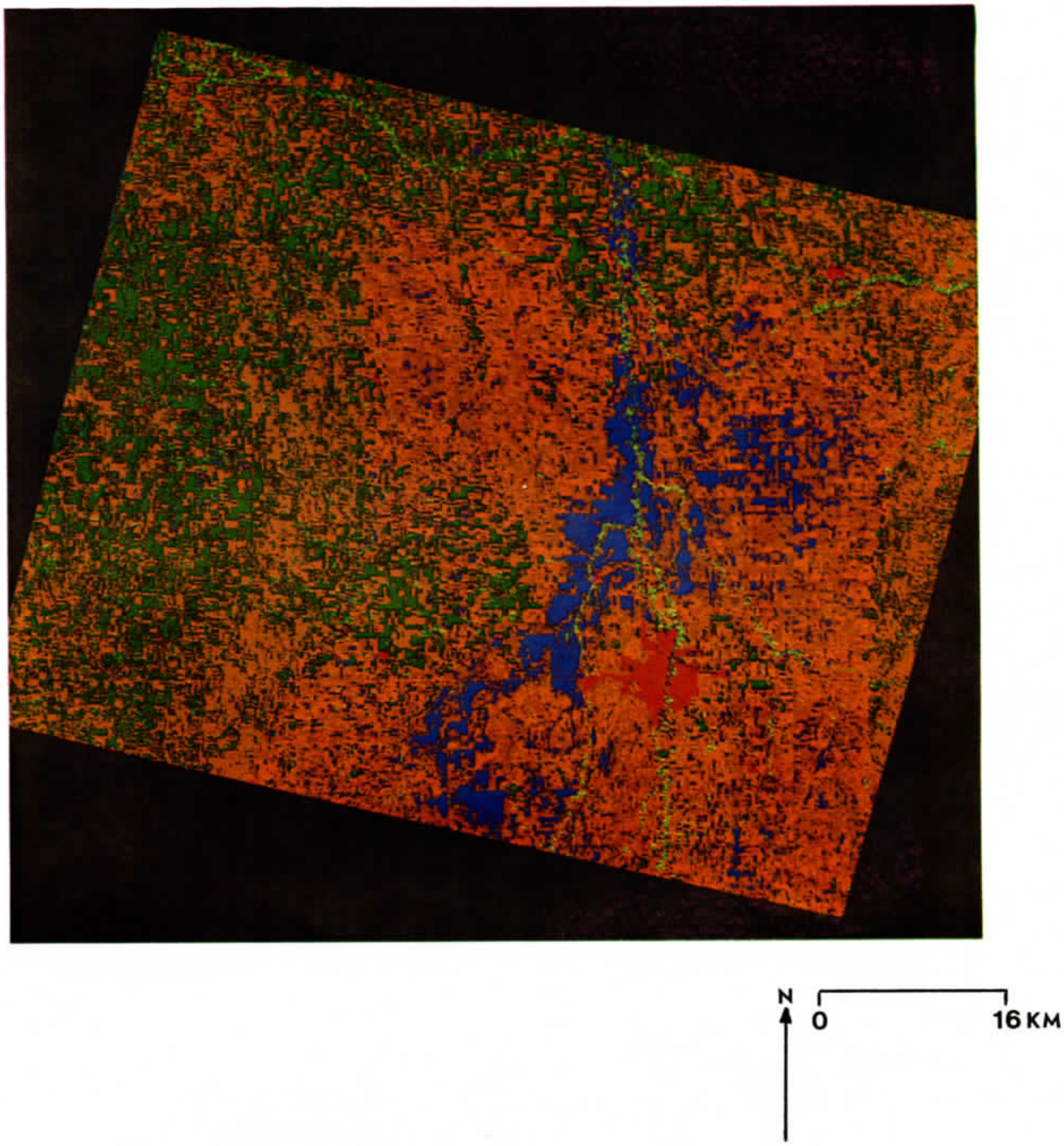

Figure 4.--Landsat classification results showing agricultural land affected by flooding. Wooded areas are shown in light green. Urban areas are shown in red, cropland not affected by flooding in dark green, and fallow fields and recently planted fields not affected by flooding are shown in light brown. Land where complete inundation occurred is shown in dark blue (standing water) and dark brown (land where there was no standing water on July 14, 1975). Land which was partly affected by flooding is shown in orange.

Table 2.--Coefficient of variation for PSU's of varying sizes. The optimum size of a PSU is determined by the size which tends to both minimize the coefficient of variation and the size of the sample unit.

Size of PSU in Pixels	Coefficient of Variation
100	.43
400	.38
1600	.33
3600	.37

Two hundred sample units, 1,600 pixels in size, were randomly allocated and the Landsat classification results were tabulated to provide first-phase estimates. Thirty second-phase sample units were randomly selected from the 200 first-phase sample units. The line and sample coordinates and latitude and longitude coordinates were computed for the four corner points of each second-phase sample unit. Each second-phase sample unit was plotted on color infrared photographs (figure 5). A dot grid with a dot density of 200 dots per square inch was used to estimate the area of cropland affected by flood conditions in each second-phase sample unit. The area estimates made on the second-phase sample units were used to develop regression coefficients to adjust estimates made in the first phase from the Landsat classification.

Using two-phase sampling statistics, 379,165 ha of agricultural land were estimated to be affected by flood conditions. It was also estimated the 148,020 ha were partially affected and 236,835 ha were completely flooded. The area estimates and associated sampling errors are shown in table 3. From the sampling errors shown in table 3, it can be seen that estimates with an accuracy of $\pm$ 10 percent at the 95 percent probability level were achieved. The small sampling errors can be attributed, in part, to a high correlation between the estimates made at each phase. When calculating sample size for this project, the correlation between the first- and second-phase measurements was estimated at 0.80. Because the correlation between measurements at each phase was higher than anticipated (0.90), the sample size could have been reduced to 155 first-phase and 7 second-phase sample units. Although this reduction in sample size would have no value to this specific project, results from this study could be used in estimating correlation between measurements at first- and second-phase for similar projects in the future. Reducing the number of sample units at each phase is significant because it would reduce the number of aerial photographs that would have to be taken and reduce the amount of photo-interpretation required to achieve a specified accuracy level.

Table 3.--Estimate of agricultural land affected by flood conditions.

Cover Class	Hectares	Sampling Error
Partial inundation	148,020	5.4%
Complete inundation	236,835	6.1%
All areas affected by flooding	379,165	3.9%

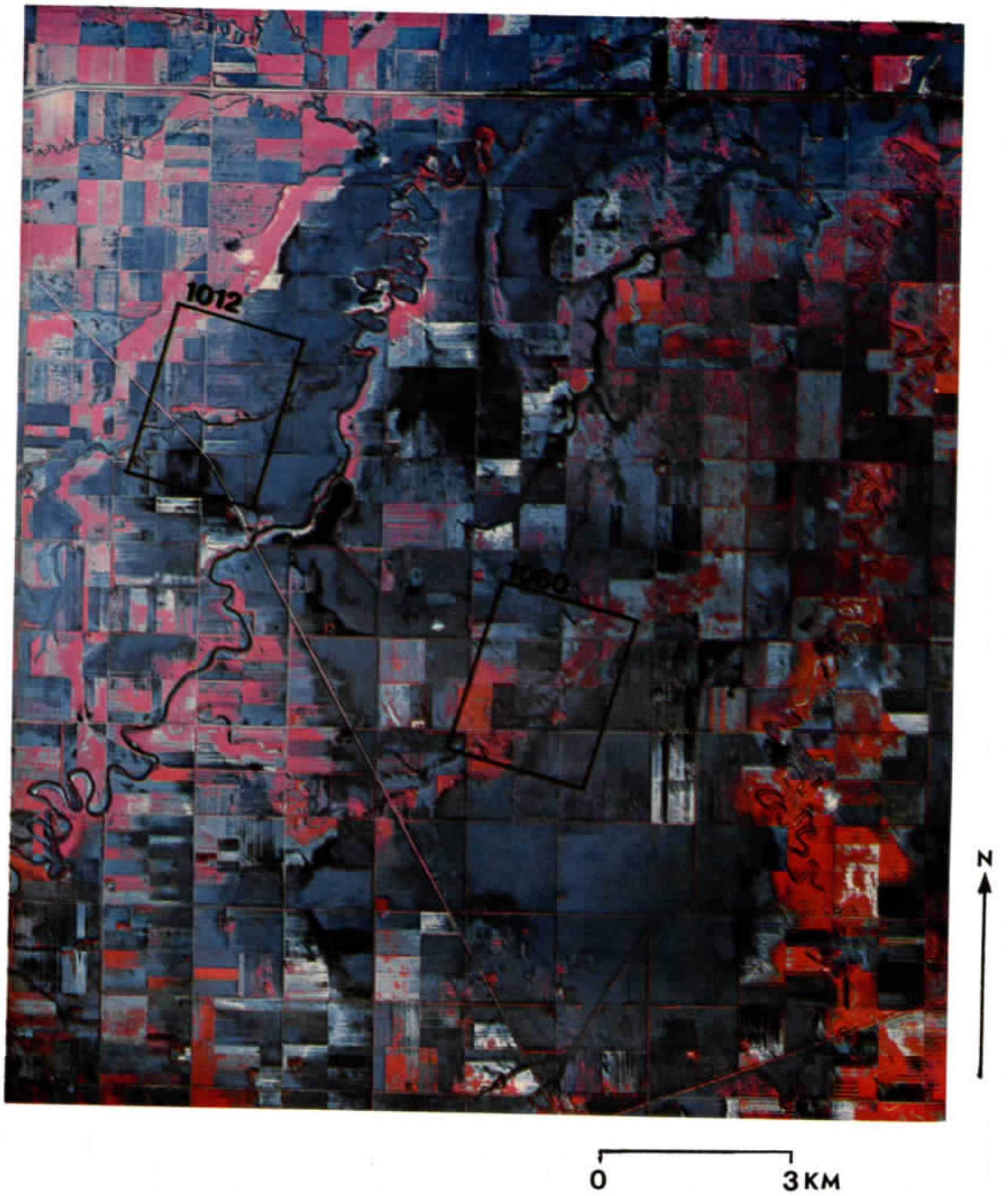

Figure 5.--Part of a NASA color infrared photograph taken on May 10, 1975, from an altitude of 18,000 m. Second phase sample plots numbered 1012 and 1060 are shown. Available ground data and a dot grid of 200 dots per square inch were used to estimate the area of agricultural land affected by flooding within each second phase sample unit.

SUMMARY

Landsat digital data and color infrared photographs were used in a multiphase sampling scheme to estimate the area of agricultural land affected by a flood. The Landsat data were classified with a maximum likelihood algorithm. Stratification of the Landsat data, prior to classification, greatly reduced misclassification errors. The classification results were used to prepare a map overlay showing the areal extent of flooding. These data also provided statistics required to estimate sample size in a two-phase sampling scheme, and provided quick, accurate estimates of areas flooded for the first phase. The measurements made in the second phase, based on ground data and photointerpretation, were used with two-phase sampling statistics to estimate the area of agricultural land affected by flooding. These results show that Landsat digital data can be used to prepare map overlays showing the extent of flooding on agricultural land and, with two-phase sampling procedures, can provide acreage estimates with sampling errors of about 5 percent.

This procedure provides a technique for rapidly assessing the areal extent of flood conditions on agricultural land and would provide a basis for designing a sampling framework to estimate the impact of flooding on crop production.

BIBLIOGRAPHY

Fleming, M.D., Berkebile, S.S., and Hoffer, R.M., 1975, Computer-aided analysis of Landsat - 1 MSS data: a comparison of three approaches, including a "modified clustering" approach: Purdue University, Laboratory for Applications of Remote Sensing, LARS Information Note 072475, 9 p.

Gialdini, M.J., Titus, S., Nichols, J., and Thomas, R.W., 1975, The integration of manual and automatic image analysis techniques with supporting ground data in a multi-stage sampling framework for timber resource inventories: three examples in Proceedings of the NASA Earth Resources Survey Symposium: Houston, Texas, National Aeronautics and Space Admin., v. 1-B, p. 1377-1388.

Heller, R.C., and Wear, J.F., 1969, Sampling forest insect epidemics with color films in Proceedings of the Sixth International Symposium on Remote Sensing of Environment: Ann Arbor, Michigan, v. II, p. 1157-1178.

Hoyer, B.E., Hallberg, G.R., and Taranik, J.V., 1974, Summary of Multispectral Flood Inundation Mapping: Iowa Geological Survey, Public Inf. Cir. 6, 57 p.

Langley, P.G., 1969, New multi-stage sampling techniques using space and aircraft imagery for forest inventory in Proceedings of the Sixth International Symposium on Remote Sensing of Environment: Ann Arbor, Michigan, University of Michigan, v. II, p. 1179-1192.

McAdams, M.P., 1976, Techniques for estimating flood produced crop damage using aerial photography in Proceedings of the Am. Soc. Photogramm. Fall Convention: Falls Church, Va., Am. Soc. of Photogramm., p. 251-266.

Meyer, William and Welch, R.I. eds., 1976, Water resources assessment *in* Reeves, R.G., ed.-in-chf., Manual of remote sensing: Falls Churck, Va., Am. Soc. Photogramm., p. 1479-1551.

Nichols, J.D., 1974, ERTS (Landsat) as an aid in timber volume inventory: Berkeley, California, University of California, Special Report to Office of Applications, NASA Headquarters 31 p.

Packer, M.R., 1976, Developing operational techniques for the use of Landsat data in identifying irrigated agriculture in Idaho, *in* Proceedings of the Am. Soc. Photogramm., Fall Convention: Falls Church, Va., Am. Soc. Photogramm., p. 267-272.